

学校编码: 10384

分类号_____密级_____

学 号: 9610005

UDC_____

学 位 论 文

犯罪统计研究

卢 海

指导教师姓名: 张兴国 副教授

申请学位级别: 硕 士

专 业 名 称: 统 计 学

论文提交日期: 1999 年 6 月

论文答辩日期: 1999 年 7 月

学位授予单位: 厦 门 大 学

学位授予日期: 1999 年 月

答辩委员会主席: _____

评 阅 人: _____

1999 年 6 月

厦门大学博硕士论文摘要库

内容提要

犯罪统计，是指国家司法机关、统计部门、有关学者定期而又系统地就犯罪现象、犯罪原因、犯罪控制所收集和出版的犯罪资料进行的统计研究。研究犯罪统计，扩大犯罪统计的影响，目的在于为某些犯罪的发生可能性以及有关嫌疑犯和被证实的犯罪情况提供合乎逻辑的准确信息。传统的犯罪统计主要是指描述犯罪现象特征和一元分析。而世上任何事物的形成、变化和发展无一不是受各种因素的影响，而且各种因素间又存在着错综复杂的关系。犯罪心理的形成和犯罪行为的发生绝非由某一因素所致，而是由一系列社会、心理、生理等多种因素相互作用或综合作用所致。因此，仅凭传统的犯罪统计分析方法已无法满足对犯罪现象进一步分析的需要。必须依靠计算机，应用现代多元统计分析方法才能研究犯罪现象中多种因素相互制约的统计规律性，进行深层次的统计分析和准确的犯罪统计预测。

犯罪多元统计分析以定性分析为经线，以定量分析为纬线，并将两者有机地联系起来，使之经纬交织，形成网络状的分析体系，具有理论方法上的深远意义。它就象一架定位器，将所有时点上的犯罪现象及其要素在立体分析的网结上予以定位分析，从定性与定量的结合上，将犯罪现象零星、游离的表现事实规引到一个崭新的、更高的层次。为我们登上精确地认识犯罪现象和有效地预防犯罪的科学殿堂打开了成功之门。

关键词：犯罪统计 回归分析 主成分分析 因子分析 聚类分析
判别分析 logistic 分析

目 录

一. 犯罪统计的历史发展	1
二. 犯罪多元回归分析	6
(一) 犯罪高峰问题研究	6
(二) 建立多元线性回归模型进行 犯罪高峰预测	8
三. 犯罪主成分分析	10
四. 犯罪因子分析	22
五. 犯罪聚类分析	30
(一) 样品聚类分析	30
(二) 变量聚类分析	34
六. 犯罪判别分析	37
七. 犯罪 logistic 分析	43
(一) logistic 分析	43
(二) 保安处分	44
(三) 家庭教育与犯罪	48

附：参考书目

犯罪统计研究

犯罪统计，是指国家司法机关、统计部门、有关学者定期而又系统地就犯罪现象、犯罪原因、犯罪控制所收集和出版的犯罪资料进行的统计研究。研究犯罪统计，扩大犯罪统计的影响，目的在于为某些犯罪的发生可能性以及有关嫌疑犯和被证实的犯罪情况提供合乎逻辑的准确信息。由于人们公认这些资料能够粗略地衡量某一犯罪辖区内或一系列具有共同的地理和人口统计学特征的辖区内的犯罪机器“严重程度”，因此是极为有益的。如果在不同的辖区之间进行比较，或者用其表示一个辖区内官方统计的犯罪历史变化，这些材料再好不过了。

一、犯罪统计的历史发展

17 世纪初，人们就已开始用统计方法来分析、说明犯罪信息，其主要代表为德国的康林（Herman Conring 1606—1682）等人，但当时对犯罪统计的理解并不系统。18 世纪末，英国人边沁（J.Bentham 1748—1832）曾设想过拟定一项与死亡统计表相类似的青少年犯罪统计表。19 世纪初，被后人认为是犯罪统计学创始人的比利时学者凯特莱（A. Quetelet 1796—1874）运用统计学的理论方法，对各种社会信息与犯罪间的联系作了大量研究，提出了犯罪的社会原因说，从而奠定了现代犯罪统计学的基础。在凯特莱的倡议下，1853 年召开了第一届国际统计学大会，犯罪统计被列为大会议程中需要收集的十类统计信息之一。继凯特莱之后，法国社会学家勒普莱（Leplay 1806—1882）及李斯特等人的研究进一步完善了犯罪统计学的体系。1827 年法国的犯罪统计年报较早地公开了官方的犯罪统计。到 19 世纪末，犯罪统计学已成为一门独立的学科。第二次世界大战以后，随着电子计算机技术的逐渐普及，运用统计方法研究犯罪问题获得迅速发展。目前许多国家都有了比较健全的犯罪统计系统和研究手段。例如，美国的各主要犯罪预防机关中，都设有至少 10 年的统计资料库，研究不同年龄组、地区、部门、职业的犯罪原因和特点，提供经验依据。在东欧，许多国家从 60 年代起就开始运用电子计算机来整理犯罪情报。日本的警视厅、最高法院和法务省每年都发行司法、检察统计年报、以及犯罪白皮书等。这些统计年报和犯罪白皮书，使得公众及时了解近年来全国犯罪的重大变化。

犯罪统计依其主体不同可分为：(1)官方犯罪统计，即司法机关和政府统计部门所进行的犯罪统计；(2)学者犯罪统计，即教学、科研部门的有关学者对犯罪及其原因所作的统计分析。依其对象不同可分为：(1)犯罪状况统计，即对相对静止时点中的犯罪密度、分布的统计；(2)犯罪动态统计，即对一段时期内犯罪变动的方向、速度的统计；(3)犯罪绝对水平统计，即用总量指标描述犯罪总规模的统计；(4)犯罪相对水平统计，即用各种相对指标描述犯罪的结构、密度（强

度)、比例、动态的统计。此外,还可分为犯罪案件统计、犯罪人统计、犯罪侵害统计。依统计内容的不同可分为:(1)描述犯罪信息特征的犯罪统计;(2)分析犯罪原因的犯罪统计;(3)预测犯罪趋势及控制效果的犯罪统计。依其形式不同可分为:(1)犯罪统计活动,包括犯罪调查统计资料的收集、整理和分析三个步骤;(2)犯罪统计资料,包括各种犯罪调查问卷、统计表格、统计年报等;(3)犯罪统计学,即运用统计学的理论和方法,分析、比较、说明犯罪现状和趋势并探讨其规律的科学。

犯罪统计是测量社会治安状况,研究犯罪手段、特征和原因,预测未来犯罪,制定刑事政策的重要方法和依据。但是,用犯罪统计结果指导刑事政策的制定是有条件的:第一,作为犯罪资料来源的已知犯罪往往少于实际中存在的犯罪,因此制定刑事政策还要考虑到犯罪黑数的大小;第二,犯罪定义在不同时空条件下可能具有不同的内容,犯罪数字在具有相同犯罪定义的时空条件下才具有可比性。因此,比较不同时空中的犯罪,还要考虑到犯罪定义是否相同。

犯罪统计资料的统计分析是指根据资料的性质、运用统计分析将资料化繁为简,互相比较,对之进行透彻的认识。统计分析不仅应有一定的体系,而且应有分析的重点和方法。犯罪统计分析的传统内容包括状况性分析、平均性分析、差异性分析、相关性分析和综合性分析等。状况性分析就是研究犯罪资料的总体轮廓,它是发掘问题的第一步,也是引起我们对问题发生兴趣的开端;平均性分析是根据统计资料平均数的性质,分析其原因,推究其结果。平均性分析使我们视野更为宽阔,分析方法更为具体,得到的结论更为正确。平均性分析的方法有算术平均、几何平均、倒数平均、中位数平均和众数平均等;差异性分析是根据统计资料差异数的性质,分析事物的差异原因或推论其结果。差异性分析重视差异状况,通过差异状况可以分析我们应研究的中心问题。表示差异的方法,可分绝对差异和相对差异,前者用原单位表示,后者用无名数表示。计算差异的方法,有标准差、平均差、四分位差等;相关性分析就是两种以上可能有关的事物,利用相关分析方法,研究彼此间的因果关系;综合性分析是根据以上四种分析所得,作出更具体的一种结论性分析。综合分析的方法可分为物理性综合分析和化学性综合分析两类,前者以不改变原有的分析内容为原则,后者改变原有的分析内容,成为一种新的分析结果。

传统的犯罪统计主要是指描述犯罪现象特征和一元分析。而世上任何事物的形成、变化和发展无一不是受各种因素的影响,而且各种因素间又存在着错综复杂的关系。犯罪心理的形成和犯罪行为的发生绝非由某一因素所致,而是由一系列社会、心理、生理等多种因素相互作用或综合作用所致。因此,仅凭传统的犯罪统计分析方法已无法满足对犯罪现象进一步分析的需要。必须依靠计算机,

应用现代多元统计分析方法才能研究犯罪现象中多种因素相互制约的统计规律性，进行深层次的统计分析和准确的犯罪统计预测。

犯罪多元统计分析是定性分析与定量分析的结合。辩证唯物主义认为，任何具体事物都既有质的规定性，又有量的规定性，是质与量的统一体。任何事物的质都是一定量的质，任何事物的量也都是一定质的量，质是量的基础，量又制约了质，没有质的量或没有量的质的事物是不存在的。这一基本原理使我们深刻认识到，定性分析与定量分析是唇齿相依，紧密联系的，是一个问题的两个方面，共处于矛盾的对立统一体中。定性分析是定量分析的基础，离开定性分析，定量分析就成了空中楼阁，如果定性分析不准，再科学的定量分析方法也毫无用武之地；另一方面，定量分析是定性分析的延伸和归宿，它为新的定性分析作出必要的准备，离开了定量分析，定性分析就会失去准确性和科学性，如果定量分析不当，定性分析就会偏离预定目标，得出错误的结论。犯罪多元统计分析以定性分析为经线，以定量分析为纬线，并将两者有机的联系起来，使之经纬交织，形成网络状的分析体系，具有理论方法上的深远意义。它就象一架定位器，将所有时点上的犯罪现象及其要素在立体分析的网结上予以定位分析，从定性与定量的结合上，将犯罪现象零星、游离的表现事实引到一个崭新的、更高的层次。为我们登上精确地认识犯罪现象和有效地预防犯罪的科学殿堂打开了成功之门。

必须指出研究犯罪统计应注意以下相关问题。首先是对犯罪黑数的认识，一个世纪以来，统计是衡量犯罪的主要尺度。而且正如人们所看到的，犯罪社会学也是随着统计学的发展及其对社会现象的应用而发展起来的，在我们对犯罪的印象方面，统计所起的作用是决定性的。犯罪统计可以引导刑事政策的制定，可以安稳舆论或使其不安，可以证明改革的正确或否定改革。但犯罪统计是十分复杂也是最困难的任务之一，各种各样的因素都对犯罪统计产生决定性影响：公众对犯罪行为的承受能力及其告发意愿、国家刑事司法系统的作用范围和有效性、犯罪统计的方法和效率等。无论如何科学的统计方法，所获得的原始数据，与生活中实际存在的犯罪数目，都不可能绝对一致，二者必然存在一定的差距。警方获悉并记录在案的，只是实际犯罪行为那座冰山露出水面的一角。尽管看起来奇怪，但我们的确对犯罪只掌握很少有用的情况。描述同样也是记数，而犯罪统计描述的只是犯罪现象最表面的东西，即可以记数的表象。法国学者罗杰·胡德在其著作《犯罪》一书中认为：没有什么比犯罪统计更能让人作出错误的解释了。诚然，犯罪统计有时更多的是让我们走入歧途，而不是给我们提供情况。因为统计所未包括的犯罪比所包括的犯罪要严重得多。官方的犯罪统计作为正式的社会主管当局的工作报告是理所当然的，但是必须认识到这种统计的局限性和弱点，并且在利用这些统计材料时，始终明白它的来历，这样才能使其成为刑事司法系统的计划和监督工作以及犯罪学研究工作的宝贵而可靠的测量工具。对所有犯罪活动和

犯罪分子，官方统计只是忠实地反映了一部分，对其他的只是片面地反映，还有的甚至没能包括。因此为了查明犯罪的规模、结构、发展和分布，必须掌握一切可以利用的统计资料，除了官方公布的犯罪统计之外，还应包括犯罪黑数的统计资料。

其次是对犯罪预测的认识。进行犯罪预测活动，虽然总是得出某种预测结果，但这并不是犯罪预测的目的，预测活动的最终目的是：依据某种预测结果，采取相应的对策，促进和保证预测结果的实现或者阻止和避免预测结果的发生。也就是说，评价犯罪预测的社会功能，主要看这种预测活动的预警作用。即对决策产生的影响作用如何，而不是也不能以这种预测结果将来能否应验为标准。事实上，对于任何一种预测结果，无论人们希望它能够应验或不应验，都离不开人的主观能动性，都要看人们能否依据这种预测结果作出正确的决策。并切实保证这种预测结果的贯彻实施，特别是对犯罪预测中得出的刺激和促进犯罪发生不利于社会安定的预测结果，我们当然不能坐望它将来应验，而应根据这些预测结果制定出预防和治理犯罪的方针政策和措施并动员全社会的力量参与社会治安的综合治理，遏制向犯罪严重化发展的趋势，以避免和阻止不利于社会安定的预测结果的出现。

二、犯罪多元回归分析

(一) 犯罪高峰问题研究

1. 历史上的五次犯罪高峰

我国建国以来，刑事案件始终处于不平衡的动态变化之中。50 年的犯罪史可归纳为五次犯罪高峰（按犯罪绝对数量）。

第一次犯罪高峰是 1950 年，当年的发案数为 51 万 3 千余件。发案率为 9.3 起/万人。这次犯罪高峰主要是旧政权和旧社会的残余分子对新政权的仇视、颠覆和破坏所造成的。以普通刑事犯罪与反革命犯罪交织为特点。

第二次犯罪高峰发生在 1961 年，发案数为 42 万余起，当年的刑事案件发案率为 6.4 起/万人，这主要是 1958 年的大跃进与人民公社化造成的灾难性后果。

第三次犯罪高峰发生在 1973 年，发案数 53 万起，当年的刑事发案率为 6 起/万人，这次犯罪高峰发生在文化大革命时期，极度混乱的社会经济和非颠倒的畸形文化道德为特点的社会大动荡直接触发了一次独特复杂的刑事犯罪高峰。

第四次犯罪高峰发生在 1981 年。当年发案数为 89 万起，发案率为 8.9 起/万人。时值改革开放的初期，文革后遗症很严重，滋生了一大批刑事犯罪分子。

第五次犯罪高峰产生于 80 年代后期。1991 年发案 236 万余起，发案率为 20 起/万人。目前还难以确定峰顶点。

2. 是否有第六次犯罪高峰， 必须进行准确的犯罪预测

由五次犯罪高峰可以看出犯罪趋势的变化与社会政治、经济形势、文化结构的变化等有密切的关系。有时这种影响是即时的，即社会政治经济形势等一旦发生变化，犯罪趋势立即随之变化。但有时犯罪趋势的变化也不一定与社会变化同时发生，如文化大革命极大地破坏了人们正常社会生活和价值观念。但当时的犯罪趋势总体变化并不显著，只是一定期后，文化结构变化，价值观念受破坏所造成的恶果才显露出来。我国第六次犯罪高峰在何时以何等规模、何种形势出现。要求我们必须作出有根据的推测和有见地的预见。

3. 传统的预测方法无法解决犯罪高峰预测问题

传统的趋势外推法和专家估计预测法虽有各自优点，却无法回答犯罪高峰等长期犯罪趋势预测问题。趋势外推法，是根据过去和现在的已构成规律

的犯罪动态数列向未来延伸的轨迹，来预测未来犯罪的趋势。这种方法是假设未来犯罪的发展趋势保持不变。但实际上犯罪趋势不可能在较长的时间内保持不变或不大变，因此趋势外推法不适于长期和中期犯罪预测。专家估计预测法，是在系统地询问社会问题专门研究人员和司法实际工作者组成的专家小组的意见基础上对结果进行科学处理和分析的预测方法。专家估计预测是一种直观预测，组织专家运用专业知识，对犯罪的过去和现在的问题进行综合分析，从中发现规律，预测未来一定时期内犯罪发展变化的趋势。诚然，专家预测法是一种较好的主观信息预测，它的最大优点是在缺乏足够统计数据和原始资料的情况下，可以作出定量估计和得到新的信息。但也正因为如此，它有不可靠的一面，因为犯罪现象极其复杂、变化多端，用专家估计预测法不可避免带有主观片面性。

4. 对犯罪趋势预测必须运用多元统计分析

在承认犯罪现象受很多犯罪影响因素制约的前提下，从中找出制约犯罪的产生、变化和发展的的重要相关因素、以其作为对犯罪趋势预测的预测因素，并通过对这些预测因素和犯罪的相关性分析，确定预测因素与犯罪的相关系数，从而测定该预测因素的预测能力，然后再根据各预测因素所起的作用大小及其变化，预测犯罪的发展趋势。通过因素分析法，我们不仅可以测知犯罪的未来趋势，而且可以根据对预测因素相关系数的分析，从预测因素中选出长期起作用的因素和暂时起作用的因素、强相关因素和阻止犯罪发生的因素等，从而在预防犯罪的过程中对症下药，为预防犯罪提供科学依据。

（二） 建立多元线性回归模型进行犯罪高峰预测

1. 回归分析模型

回归分析是研究一个变量（通常称为因变量或反应变量）与一组变量（通常称为自变量或解释变量）之间关系的重要手段。变量之间的这种关系通常可用方程式来表示：

$$Y_i = b_0 + b_1 X_{i1} + b_2 X_{i2} + \cdots + b_q X_{iq} + e_i$$

式中： Y_i 为第 i 个个体因变量的值

b_0 为常数项

b_i 为偏回归系数

e_i 为误差项

$X_{i1}, X_{i2}, \cdots, X_{iq}$ 为第 i 个个体 q 个自变量的值。

回归分析中要解决的问题有：

- a. 建立回归模型 — 包括模型选择，并进行显著性检验；

- b. 进行回归诊断;
- c. 回归预测。

2. 建立犯罪高峰预测模型应注意的问题

在对犯罪趋势进行预测的过程中, 必须全面把握制约犯罪的各种因素, 既要重视与犯罪相关性高的因素, 也要注意与犯罪相关性不高的因素; 既要重视与犯罪存在正相关性—即可能诱发、促使犯罪的因素, 也要注意与犯罪存在负相关性—即可能抑制、阻碍犯罪发生的因素, 并且在掌握这些因素的内部机理的基础上, 综合考虑这些因素与犯罪的相关性, 切忌犯罪预测和因素评估的片面性。但是在现代社会, 犯罪现象是一个非常复杂的社会现象, 它与众多的社会环境因素和犯罪者自身条件有着程度不同, 直接或间接的联系。这种联系有广泛性、深刻性、复杂性、易变性, 许多机理我们至今还不甚了解, 甚至完全不了解。这就使得我们在进行犯罪预测时, 事实上不可能把制约犯罪消长和变化的各种因素都全面考虑到。因此, 在目前的条件下, 我们进行犯罪趋势预测, 只能要求把握那些对未来犯罪发展动态和趋势起主要、直接制约作用的因素, 通过对这些因素的相关性及其发展趋势的深入调查和分析来对犯罪发展的一般趋势进行评估和推断。

因此在建立犯罪高峰预测模型时, 既要防止多重共线性, 又要防止遗漏重要变量。遗漏重要变量将使参数估计值产生偏倚, 因此在设定模型时应尽量避免这种情况发生, 即使是资料限制或为克服多重共线性不得不舍弃某些重要变量, 也应当了解偏倚的情况, 从而在应用估计结果时审慎考虑这种误差可能带来的影响。引入不相关的解释变量虽不会导致参数估计值的偏倚, 但也可能把一些无法判定是否合适的解释变量统统引入模型, 这样一来, 不仅增大了参数估计值的方差, 而且容易导致多重共线性。因为引入的变量之间往往是相关的或与模型中的其他解释变量相关, 虽然单一回归系数的显著性检验和若干回归系数的显著性检验能把不相关的解释变量拒于模型之外, 但如果引入的变量太多, 多重共线性确已发生, 就无法分辨出各个解释变量对因变量的影响作用, 这些检验也将失败。

3. 变量逐步引进法建立犯罪高峰预测模型

一种既不遗漏重要变量又可避免滥引进的方法是把所有可能的解释变量(即影响犯罪的所有可能因素)罗列出来, 根据犯罪理论知识, 选择几个重要变量建立一个基本模型:

$$Y = b_0 + b_1 x_1 + \cdots + b_q x_q + u \quad (1)$$

得到的 OLS 估计式为:

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 x_1 + \cdots + \hat{b}_q x_q \quad (2)$$

该式的残差平方和为 $e_1'e_1$ ，然后在 (1) 式中引入新的解释变量 X_{q+1} ，得：

$$Y = b_0 + b_1 x_1 + \cdots + b_{q+1} X_{q+1} + \cdots + b_k x_k + e_i$$

将此式的 OLS 残差记为 e_2 ，则：

$$F = \frac{(e_1'e_1 - e_2'e_2) / (k-q)}{e_2'e_2 / (k-q)} \sim F(k-q, n-k)$$

在显著性水平 α 上，若 $F > F_{\alpha}(k-q, n-k)$ ，则认为新的解释变量 X_{q+1} ， \cdots ， X_k 显著地提高了原模型的拟合程度，是被遗漏的重要解释变量。

若 $F < F_{\alpha}(k-q, n-k)$ ，则认为 X_{q+1} ， \cdots ， X_k 是不相关的解释变量。通过以上逐步引进和排除，则犯罪模型建立。

三．犯罪主成分分析

犯罪研究所涉及的问题往往比较复杂，这是因为影响犯罪的因素较多，需要考察的变量也比较多。例如对于抢劫、杀人等犯罪行为，其动机是多种多样的，收集的资料中包含的信息也是各种各样的。然而重叠的、低质量的信息越多，越不利于我们对犯罪现象作出分析判断。在大部分实际问题中，变量之间具有一定的相关性，人们自然希望找到较少的几个彼此不相关的综合指标尽可能多地反映原来的多变量的信息。

主成分分析（principal component analysis）方法正是实现这一目的的有效途径之一。主成分简而言之就是由原变量 $X_1 \cdots \cdots X_m$ 线性组合的互不相关，且未丢失任何信息的新变量，也称为综合变量。多指标的主成分分析常被用来寻找某种现象或事物的综合指标，并给综合指标所蕴藏的信息以恰当解释，以便深刻地揭示事物内在的规律。

主成分分析实例

下面的数据是美国第 1—50 个州中每 100,000 个人中 7 种犯罪的人数数据，七个变量的众多数据使我们无法评价各州的治安和犯罪情况。

	y1	y2	y3	y4	y5	y6	y7
	凶杀	强奸	抢劫	袭击	盗窃	偷窃	盗车
No	murde	rape	robber	assault	burglary	larceny	Auto
1	14.2	25.20	96.80	278.30	1135.30	1881.90	280.70
2	10.80	51.60	96.80	284.00	1331.70	3369.80	753.30
3	9.50	34.20	138.20	312.30	2346.10	4467.40	439.50
4	8.80	27.60	83.20	203.40	972.60	1862.10	183.40
5	11.50	49.40	287.00	358.00	2139.40	3499.80	663.50
6	6.30	42.00	170.70	292.90	1935.20	3903.20	477.10
7	4.20	16.80	129.50	131.80	1346.00	2620.70	593.20
8	6.00	24.90	157.00	194.20	1682.60	3678.40	467.00

续上表

9	10.20	39.60	187.90	449.10	1859.90	3840.50	351.40
10	14.70	31.30	140.50	256.50	1351.10	2170.20	297.90
11	7.20	25.50	128.00	64.10	1911.50	3920.40	489.40
12	5.50	19.40	39.60	172.50	1050.80	2599.60	237.60
13	9.90	21.80	211.30	209.00	1085.00	2828.50	528.60
14	7.40	26.50	123.20	153.50	1086.20	2498.70	377.40

15	2.30	10.60	41.20	89.30	812.50	2685.10	219.90
16	6.60	22.00	100.70	180.50	1270.40	2739.30	244.30
17	10.10	19.10	81.10	123.30	872.20	1662.10	245.40
18	15.50	30.90	142.90	335.50	1165.50	2469.90	337.70
19	2.40	13.50	38.70	170.00	1253.10	2350.70	246.90
20	8.00	34.80	292.10	358.90	1400.00	3177.70	428.50
21	3.10	20.80	169.10	231.60	1532.20	2311.30	114.00
22	9.30	38.90	261.90	274.60	1522.70	3159.00	545.50
23	2.70	19.50	85.90	85.80	1134.70	2559.30	343.10
24	14.30	19.60	65.70	189.10	915.60	1239.90	144.40
25	9.60	28.30	189.00	233.50	1318.30	2424.20	378.40
26	5.40	16.70	39.20	156.80	804.90	2773.20	309.20
27	3.90	18.10	64.70	112.70	760.00	2316.10	249.10
28	15.80	49.10	323.10	355.00	2453.10	2412.60	559.20
29	3.20	10.70	23.20	76.00	1041.70	2343.90	293.40
30	5.60	21.00	180.40	185.10	1435.80	2774.50	511.50
31	8.80	39.10	109.60	343.40	1418.70	3008.60	259.50
32	10.70	29.40	427.60	319.10	1728.00	2782.00	745.80
33	10.60	17.00	61.30	318.30	1154.10	2037.80	192.10
34	.90	9.00	13.30	43.80	446.10	1843.00	144.70
35	7.80	27.30	190.50	181.10	1216.00	2696.80	400.40
36	8.60	29.20	73.80	205.00	1288.20	2228.10	326.80
37	4.90	39.30	124.10	286.90	1636.40	3506.10	388.90
38	5.60	19.00	130.30	128.00	877.50	1624.10	333.20
39	3.60	10.50	86.50	201.00	1489.50	2844.10	791.40
40	11.90	33.00	105.90	485.30	1613.60	2342.40	245.10

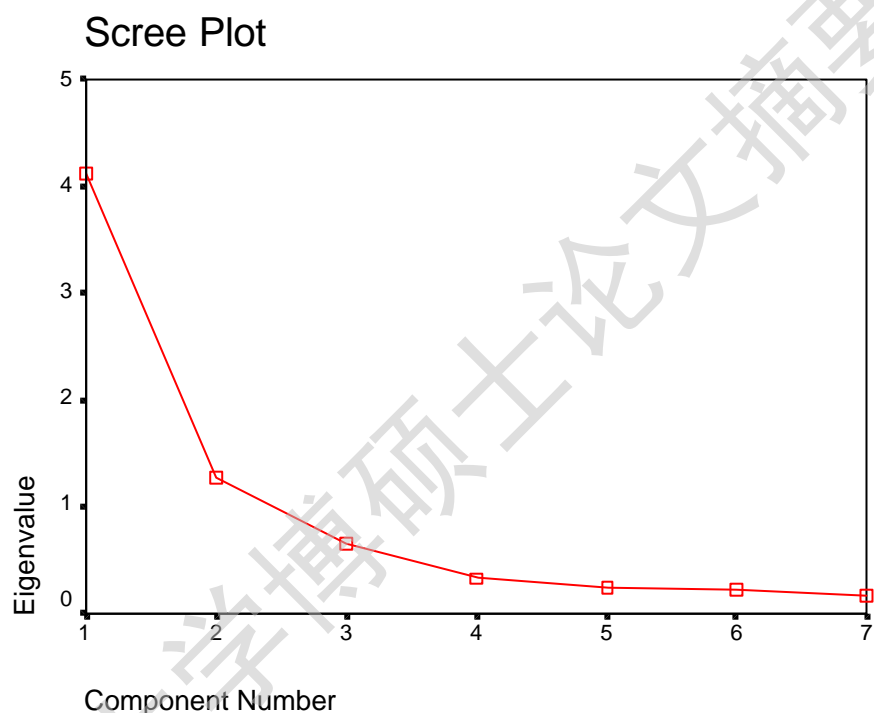
续上表

41	2.00	13.50	17.90	155.70	570.50	1704.40	147.50
42	10.10	29.70	145.80	203.90	1259.70	1776.50	314.00
43	13.30	33.80	152.40	208.20	1603.10	2988.70	397.60
44	3.50	20.30	68.80	147.30	1171.60	3004.60	334.50
45	1.40	15.90	30.80	101.20	1348.20	2201.00	265.20
46	9.00	23.30	92.10	165.70	986.20	2521.20	226.70
47	4.30	39.60	106.20	224.80	1605.60	3386.90	360.30

48	6.00	13.20	42.20	90.90	597.40	1341.70	163.30
49	2.80	12.90	52.20	63.70	846.90	2614.20	220.70
50	5.40	21.90	39.70	173.90	811.60	2772.20	282.

现根据上表数据进行主成分分析，主要结果如下：

表（1）特征值碎石图



上图中 scree plot 为特征值碎石图

Eigenvalue 为主成份特征值

Component number 为主成分个数

从表（1）中可以看出主成分特征值共有 7 个，在第二个特征值处有急转弯，第三个及其以后的主成分特征值都较低(<1)，因此保留前两个主成分是合理的。

表(2)主成分特征值贡献率

Total Variance Explained

Component	Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %
1	4.123	58.903	58.903
2	1.269	18.128	77.031

Extraction Method: Principal Component Analysis

上图中 total variance explained 为主成分所解释方差

component 为主成分

total 为主成分特征值

% of variance 为主成分特征值贡献率

cumulative %为主成分特征值贡献率累计

extraction method 为提取方法

principal component analysis 为主成分分析

从表(2)中可以看出前两个主成分特征值所解释的方差累计为 77%。因此,可以说前两个主成分已经概括了大部分信息。

表(3) 主成分矩阵

Component Matrix^a

	Component	
	1	2
Y1	.610	-.698
Y2	.881	-.145
Y3	.814	-3.85E-02
Y4	.797	-.376
Y5	.879	.207
Y6	.646	.628
Y7	.699	.424

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

上图中 component 为主成分

component matrix 为主成分矩阵

extraction method 为提取方法

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库